

Introducción a ML y GenAI

Evaluación y validación de modelos

Ariel Ramos Vela

10-10-2024

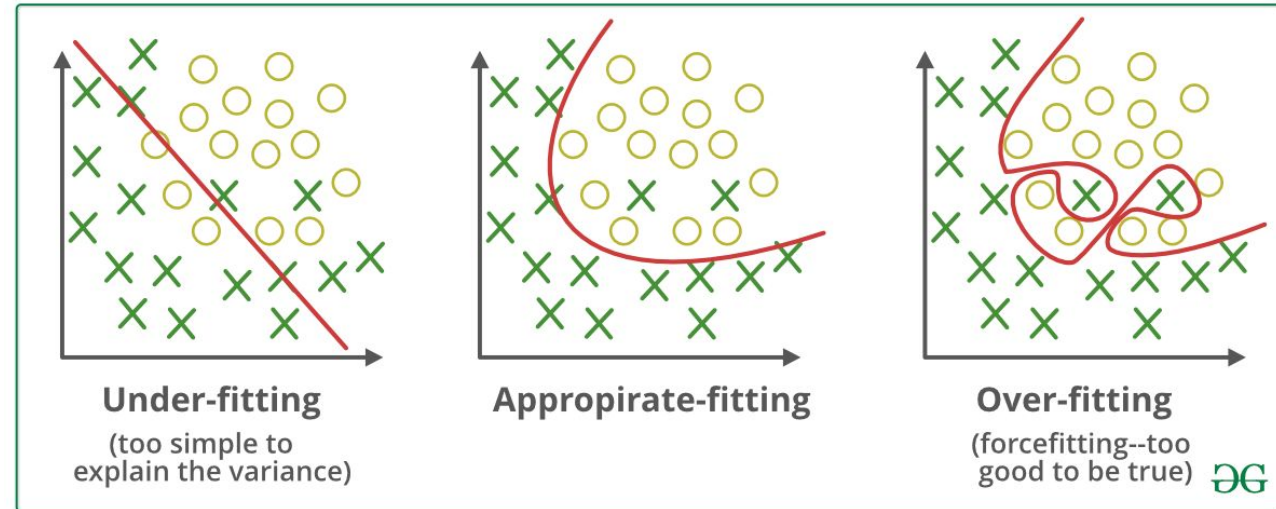
Agenda

1. Sobreajuste y Subajuste
2. Validación Cruzada
3. Ejemplo práctico (Python)
4. Métricas de Rendimiento
5. Proyecto

Sección I: Sobreajuste y Subajuste

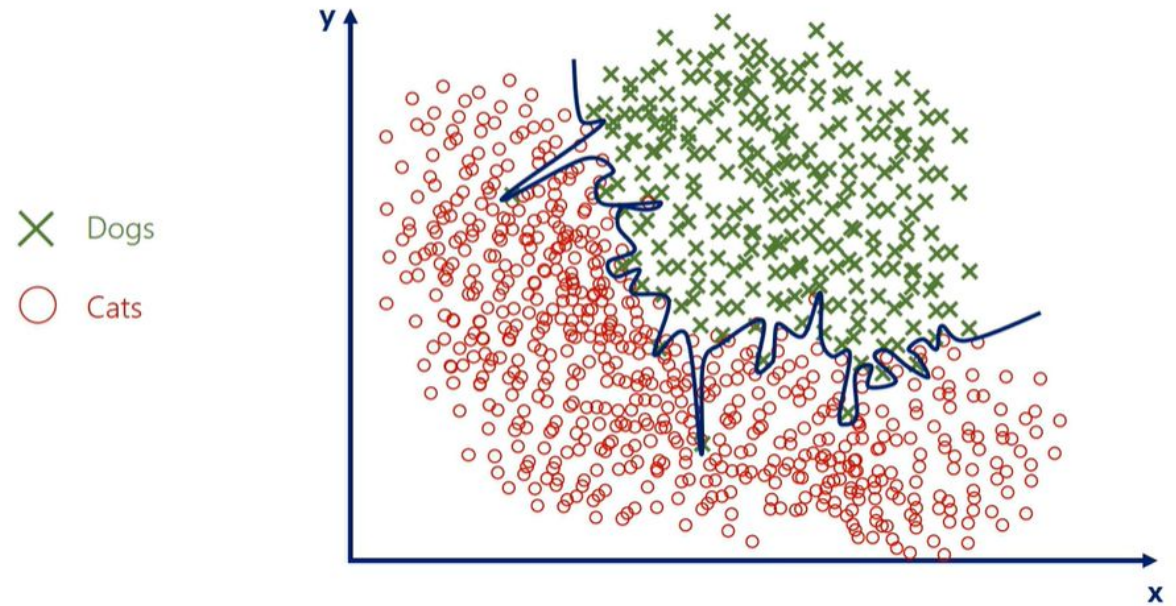
Sobreajuste (Overfitting): Modelo muy complejo que se ajusta demasiado bien al conjunto de entrenamiento, pero falla en generalizar a datos nuevos.

Subajuste (Underfitting): Modelo demasiado simple que no captura patrones importantes de los datos.



Factores que Conducen a Sobreajuste y Subajuste

- **Sobreajuste:**
 - Modelo demasiado complejo
 - Exceso de características (features)
 - Datos insuficientes
- **Subajuste:**
 - Modelo demasiado simple
 - No suficiente entrenamiento



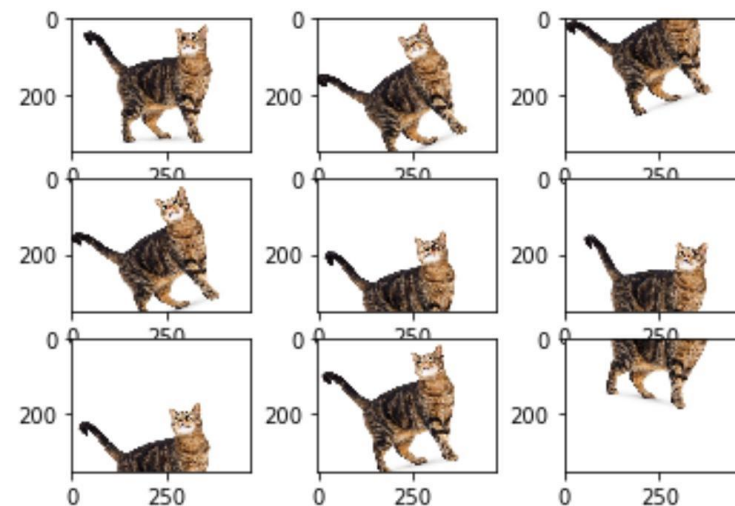
¿Sobreajuste o subajuste?

Cómo Evitar el Sobreajuste

- **Regularización (L1, L2)**
- **Simplificación del Modelo**
(Reducir complejidad)
- **Recolección de más datos**
- **Data Augmentation** (para imágenes y texto)

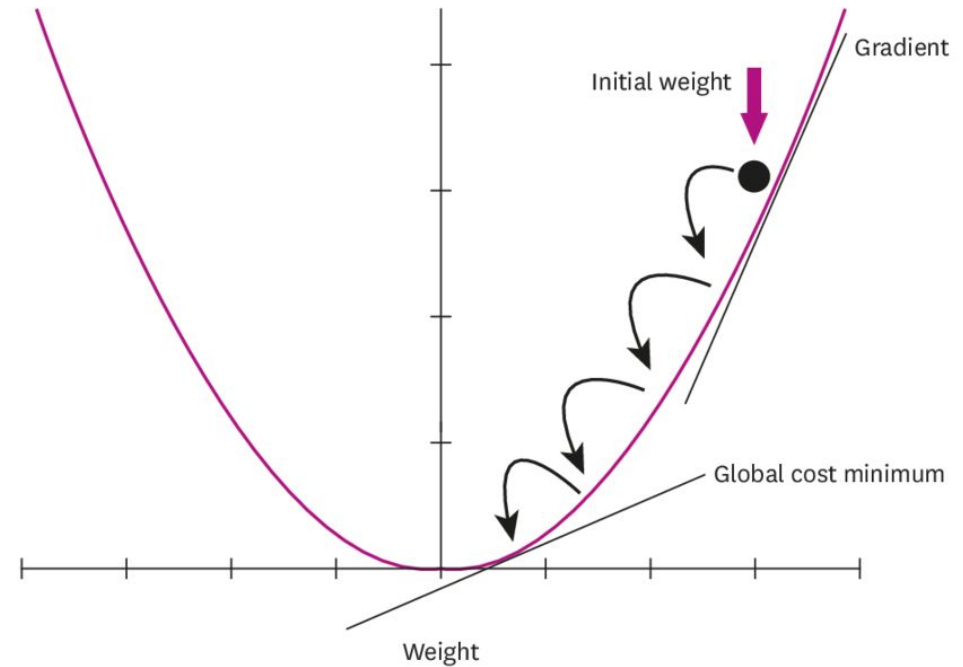
$$J = \text{Error} + \lambda \sum |w_i|$$

$$J = \text{Error} + \lambda \sum w_i^2$$



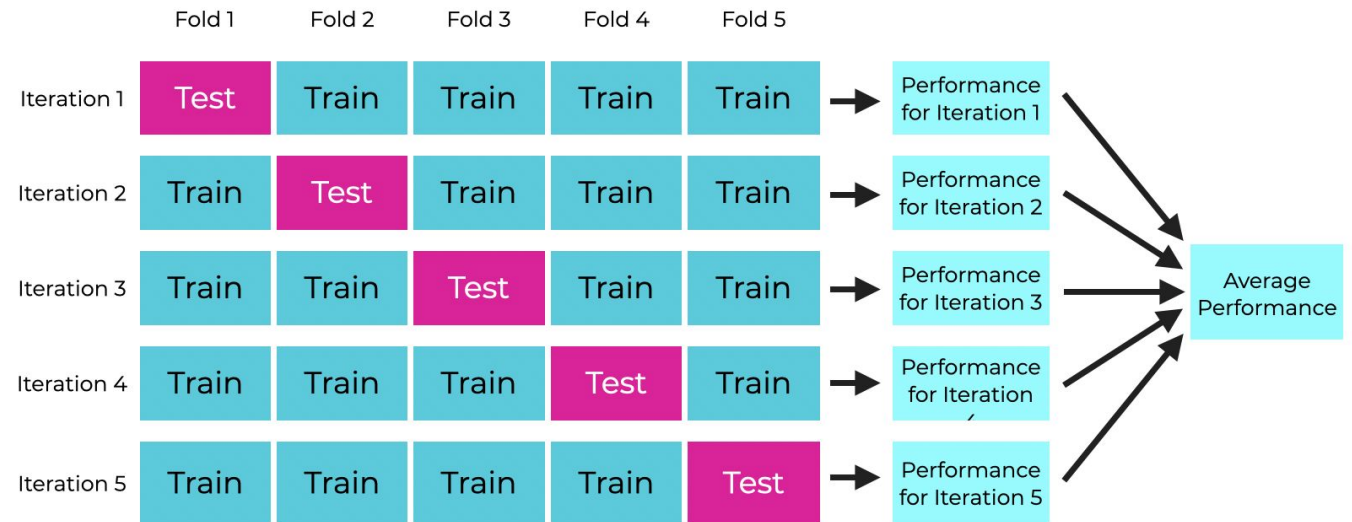
Cómo Evitar el Subajuste

- **Modelos más complejos**
(aumentar capacidad)
- **Mayor entrenamiento** (más tiempo, mejores hiperparámetros)
- **Añadir características relevantes**



Sección II: Validación Cruzada (Cross Validation)

Dividir los datos en varios subconjuntos (folds) y evaluar el modelo varias veces para evitar la dependencia en un único conjunto de validación.



K-Fold Cross Validation

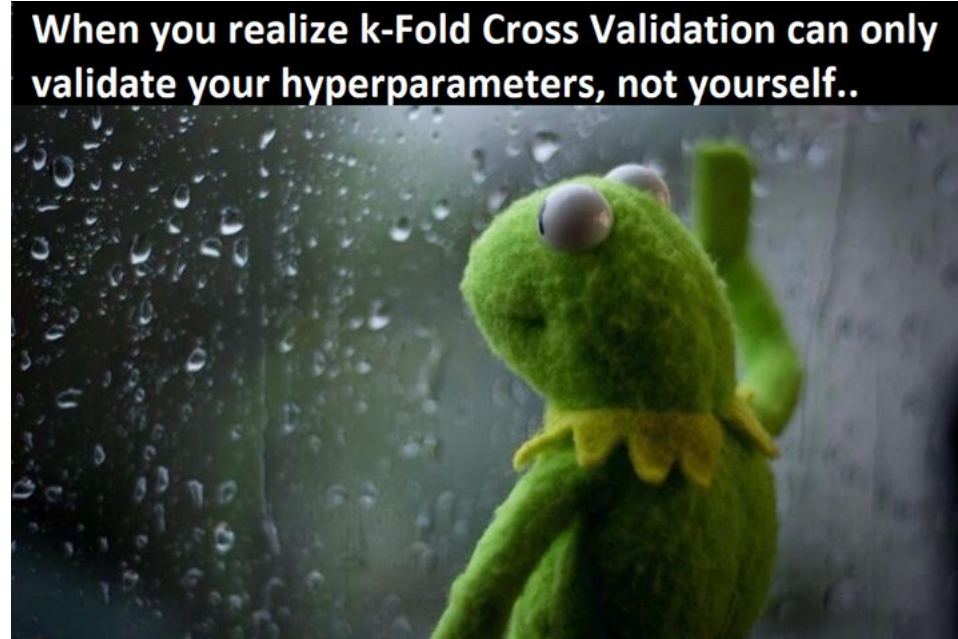
Dividir el conjunto de datos en K partes, entrenar en K-1 y validar en el restante, repitiendo K veces.

4-fold validation (k=4)



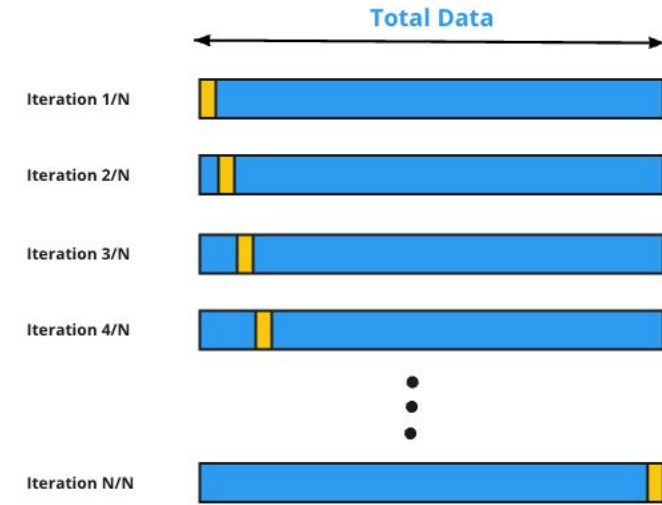
Beneficios de la Validación Cruzada

- **Reducir la varianza en la estimación del rendimiento**
- **Aprovechamiento de los datos:**
Cada punto de datos se usa para entrenar y validar.
- Evita sesgos de entrenamiento y selección errónea de hiperparámetros.



Tipos de Validación Cruzad

- K-Fold Cross Validation
- Leave-One-Out Validation (LOOCV) Cross
- Stratified K-Fold (para datos desbalanceados)



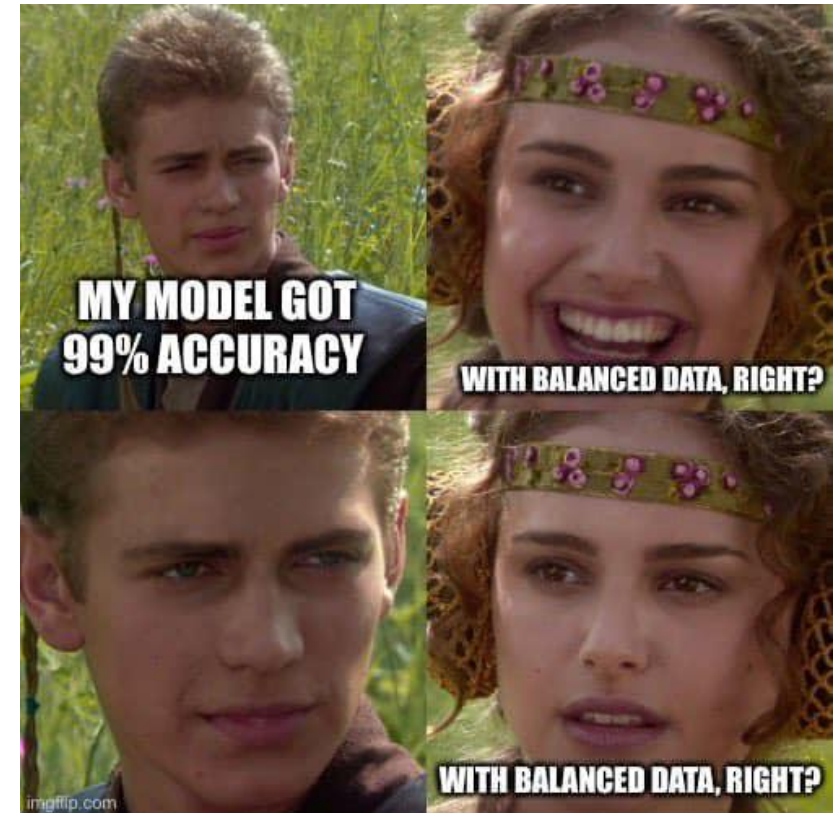
Cuándo Usar Validación Cruzada

- **Datasets pequeños:** Permite aprovechar al máximo los datos.
- **Elección de hiperparámetros:** Ideal para ajustar parámetros sin sobreajuste.
- Limitaciones:
 - **Costo computacional alto** para grandes datasets.
- Ejemplo Práctico (Python)

Sección III: Métricas de Rendimiento






¿Por qué son importantes las métricas de rendimiento?

- Diferentes métricas capturan diferentes aspectos del desempeño del modelo.
- Ayudan a evitar el enfoque exclusivo en la precisión o en métricas que no reflejan la realidad.



Métricas de Clasificación

- Exactitud (Accuracy)
- Precisión (Precision)
- Recall (Sensibilidad)
- F1-Score
- Matriz de Confusión

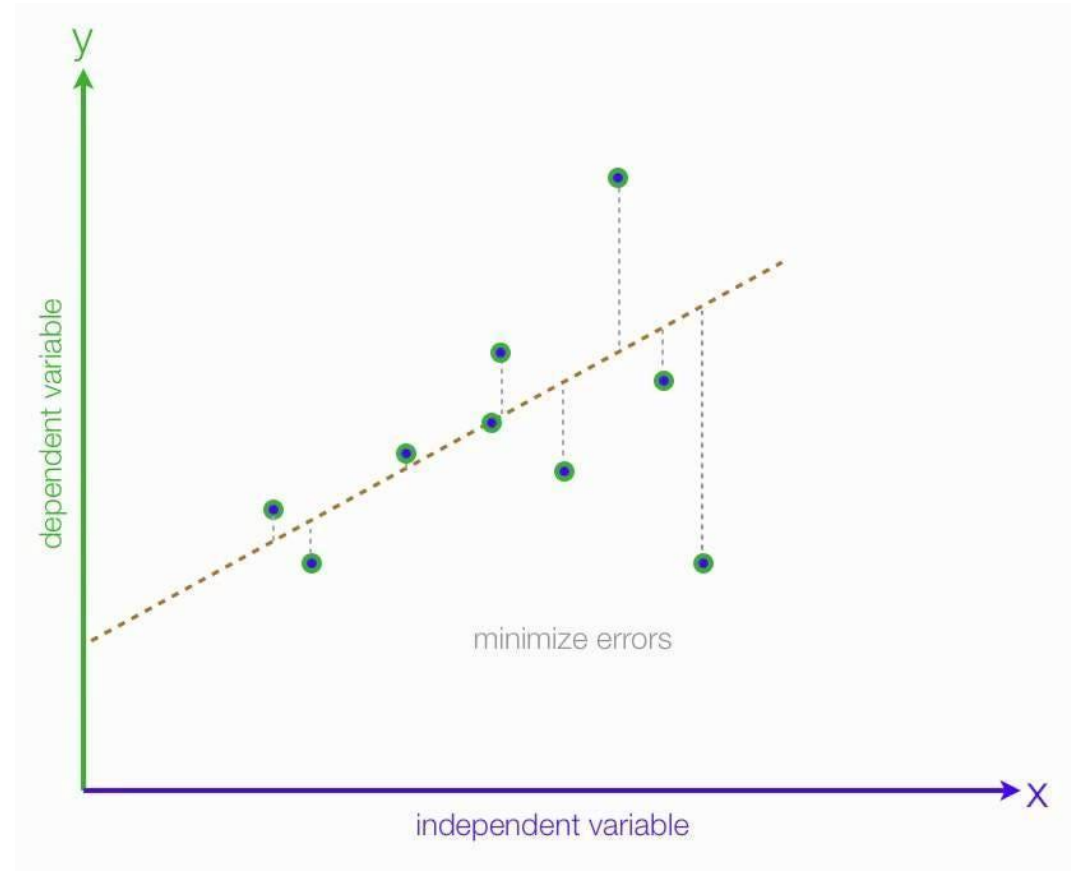
	Predicted	
	Animal	Not animal
	Animal	Not animal
Actual	  	
	 	

True Positives	2
True Negatives	3
False Positives	0
False Negatives	1

Accuracy	83%	$\frac{3+2}{3+2+0+1}$
Precision	75%	$\frac{3}{3+1}$
Recall	100%	$\frac{3}{3+0}$
F1 score	86%	$2 \cdot \frac{0.75 \cdot 1}{0.75 + 1}$

Métricas de Regresión

- **MSE (Mean Squared Error)**
- **MAE (Mean Absolute Error)**
- **R^2 (Coeficiente de Determinación)**



Conclusiones

- Importancia de balancear entre sobreajuste y subajuste.
- Uso de validación cruzada para generalización efectiva.
- Métricas de rendimiento como herramienta clave para la evaluación del modelo.